

**AMENDMENT**

- 1 1. (Original) A method for creating a cleansed output file from an input file, comprising the  
2 steps of:
- 3 (a) selecting an input file containing a plurality of data records;  
4 (b) selecting a reference files, said reference file containing a plurality of data  
5 records;  
6 (c) computing a search key; and  
7 (d) for each said data record in said input file:
- 8 (i) retrieving said data record from said input file on remote storage;  
9 (ii) searching said reference file with a matcher process for all said data  
10 records in said reference file that match said search key and reading each said data  
11 record from said reference file that matches said search key, thereby generating a  
12 candidate data record list;  
13 (iii) searching said candidate data record list and determining a matching data  
14 record, wherein said matching data record matches said data record in said input  
15 file;  
16 (iv) creating a new cleansed data record;  
17 (v) cleansing said data record of said input file according to said matching  
18 data record, thereby generating verified information;  
19 (vi) writing said verified information into said new cleansed data record; and  
20 (vii) writing said new cleansed data record to a cleansed output file;
- 21 wherein said steps (d)(i) through (d)(vii) are performed in a single pass through said data  
22 records of said input file and in a single pass through said reference file, such that each  
23 data record of said input file is read from a remote storage location only once, each said  
24 matching data record of said reference file is read from a remote storage location only  
25 once, and each said new data record to said cleansed output file is written to a remote  
26 storage location only once.

- 1 2. (Original) The method according to claim 1, further comprising two or more reference  
2 files and further comprising the step of:  
3 (e) repeating said step (d)(v) for said data record of said input file wherein said  
4 matcher cleanses said data record of said input file using said two or more reference files.
- 1 3. (Original) The method according to claim 1, further comprising two or more reference  
2 files and two or more matcher processes, wherein each said matcher process access one  
3 of said one or more reference files and said data record of said input file resides in local  
4 memory while being processed by each of said matcher processes and each of said  
5 reference files.
- 1 4. (Original) The method according to claim 3, further comprising the step of:  
2 (e) recycling one said data record of said input file through one or more matcher  
3 processes while said data record of said input file resides in local memory, wherein said  
4 recycling processes said data record of said input file through one or more of said one or  
5 more reference files previously accessed.
- 1 5. (Original) The method according to claim 3, further comprising the steps of:  
2 (e) each said matcher process determining whether a subsequent matcher process  
3 should process said data record of said input file, wherein said determining results in no  
4 additional processing is performed on said data record of said input file, one or more said  
5 matcher processes are skipped, or a subsequent matcher process is changed.
- 1 6. (Original) The method according to claim 3, further comprising two or more search keys  
2 wherein each said matcher process accesses one said search key.
- 1 7. (Original) The method according to claim 2, further comprising two or more search keys,  
2 wherein said matcher process performs step (d)(ii) using each of said search keys.

- 1 8. (Original) The method according to claim 1, wherein said search key comprises: a  
2 predefined number of digits of the USPS ZIP+4 Code, a predefined number of digits  
3 representing an address number of a postal patron, a predefined number of alphanumeric  
4 characters representing a street name of the postal patron, and a predefined number of  
5 alphabetic characters representing the postal patron.
- 1 9. (Original) The method according to claim 1, wherein said reference file is organized such  
2 that two or more data records of said reference file that match a predefined search key are  
3 stored on remote storage in a primary physical block of memory, such that said matcher  
4 process reads said candidate data record list in said step (d)(ii) with a single memory read  
5 command.
- 1 10. (Original) The method according to claim 9, wherein said reference file is organized such  
2 that two or more data records of said reference file that match a said search key are stored  
3 on remote storage in a primary physical block and an overflow physical block of  
4 memory, such that said matcher process reads said candidate data record list in said step  
5 (d)(ii) with two memory read commands.
- 1 11. (Original) The method according to claim 1, wherein said determining said matching  
2 record by said matcher process of said step (d)(iii), comprises the steps of:  
3 (1) for each data field in said search key, determining a degree of match between said  
4 data record of said input file and a data record of said reference file in said  
5 candidate data record list, said degree of similarity ranging in value from an  
6 identical value to no identifiable similarity;  
7 (2) arranging said degree of match determined in said step (d)(iii)(1) in a table;  
8 (3) determining whether each row of table in said step (d)(iii)(2) represents a match  
9 or no match of said data record of said input file and said data record of said  
10 candidate data record list; and

11           (4)     establishing for each row of said table in said step (d)(iii)(2) a match or no-match  
12                     value.

1    12.    (Original) The method according to claim 1, wherein said data records of said input file  
2           contain a mixture of personal and business data records, said business records containing  
3           zero or more contact names and addresses, and said search key containing data elements  
4           selected from the group consisting of personal name and demographical information.

1    13.    (Original) The method according to claim 1, wherein in said step (d)(iii) said matcher  
2           process matches a female's maiden name to her married name, comprising the steps of:

3           (1)     matching a confirming piece of information, such as all or part of the: Social  
4                     Security Number (SSN), Date of Birth (DOB), Driver's License State and  
5                     Number, Phone Number, USPS Delivery Point Code, or other identifying  
6                     information of the female in said matching data record and said data record of  
7                     said input file, or matching a first name and at least a middle initial in said  
8                     matching data record and said data record of said input file;

9           (2)     matching a portion of a last name in said matching data record and said data  
10                    record of said input file; and

11          (3)     confirming that the gender of the female is selected from the group consisting of  
12                    female, indeterminate, or blank.

1    14.    (Original) The method according to claim 1, further comprising the steps of:

2           (e)     generating a unique front door key for a front door of an address wherein said  
3           unique front door key comprises a USPS ZIP+4 Code + a predefined number of digits of  
4           a house number and a predefined number of characters of a unit number;

5           (f)     assigning said unique front door key to said cleansed data record of said output  
6                    file; and

7           (g)     maintaining a reference database that correlates all new DPC assignments for an  
8                    address to said unique front door key.

1 15. (Cancelled)

1 16. (Original) The method according to claim 1, wherein the method provides data to a third  
2 party with associated software to allow use of the provided data, without allowing the  
3 third party direct access to the data in clear text form, wherein said reference database is  
4 un-encrypted and said method further comprises:

5 (e) a means for encrypting said reference file and extracting a key to allow  
6 aggregating records to be considered for match between said input file and said reference  
7 database;

8 (f) a means for encrypting said input file and extracting said key for said reference  
9 database;

10 (g) a means for comparing said encrypted input file against said encrypted reference  
11 file, and if there is an acceptable match in encrypted form to cause said encrypted  
12 reference data record to be unencrypted, and the required data from said reference file be  
13 appended to said data record of said input file; and

14 (h) reporting the data content used for royalty reporting purposes.

1 17. (Original) The method according to claim 16, wherein the method prevents the third party  
2 which has been provided said encrypted reference file and associated software for use of  
3 said encrypted reference file from modifying usage reporting of said encrypted reference  
4 file without detection of such modifications.

1 18. (Original) The method of claim 1, wherein said method provides USPS postal coding  
2 without depending on the ZIP Code or Last Line City and State Abbreviation to be  
3 correct, comprising the steps of

4 (e) wherein said reference file contains USPS ZIP+4 data, indexing said reference file  
5 on:

6 (i) Delivery Point Code (DPC);

- 7 (ii) Five Digit ZIP Code+Right most N (such as five) characters of House  
8 Number+First M (such as three) characters of street name;  
9 (iii) House number+SOUNDEX street name (without prefixes and  
10 suffixes)+unit number (if any)+Street PreDirectional (if any)+Street Post  
11 Directional (if any)+Street Type (if any); and  
12 (iv) State+City Name;

13 wherein said matcher process comprises a means to retrieve on any of the available keys  
14 using information from an input record and then to match, using an available matching  
15 tool that input record to the base record to then assign the postal codes and parsed address  
16 text, and a means for appending a footnote code to each input record that describes the  
17 degree of match to each data field of the input record.

- 1 19. (Original) The method according to claim 3, wherein the method improves match rates  
2 for historic and "dirty data" address files, wherein one said reference file is constructed  
3 from the Federal Information Processing Standards (FIPS) Named Populated Places file,  
4 said search key indexes on place name and state, and said matcher process returns a ZIP  
5 Code from such index Named Populated Places reference file, further comprising:  
6 (e) matching using this ZIP where there was no ZIP Code present in the input record  
7 or the ZIP Code was different for the associated place name, or no match was obtained  
8 with the input record provided ZIP Code; comprising;  
9 i. inputting the original input ZIP Code to obtain a current ZIP Code;  
10 ii. inputting the original state and place name to obtain the current ZIP Code;  
11 or  
12 iii. proceeding with a matching process using the new ZIP Code obtained;  
13 (f) creating a second reference file constructed from the long term history of  
14 discontinued ZIPs and associated discontinued place names and the new ZIP Code and  
15 new place name;  
16 (g) indexing said second reference file on the state and place name and also on ZIP  
17 Code;

Serial No. 10/820,375

Attorney Docket No.: 070375.00003

Amendment and Response to Restriction Requirement

- 18           (h)     proceeding, if no match was found with the matching process using either of the  
19           following:
- 20                 (i)     inputting the original input ZIP Code to obtain a current ZIP Code, or  
21                 (ii)    inputting the original state and place name to obtain the current ZIP Code;  
22                 and  
23           (i)     proceeding with a matching process using the new ZIP Code obtained.